Appendix B. Univariate Power Analysis for Trend in Basal Area and Sapling Density

Kathryn M. Irvine

Abstract.

This report summarizes the preliminary univariate power analysis for the Klamath Network Vegetation protocol. The power analysis is based on the pilot study data collected at Crater Lake and US Forest Service Forest Inventory and Analysis (FIA) data from the Oregon Cascades. Pilot data are from one sample period and FIA data are from two periods that were 1 year apart. The variables analyzed were total tree basal area and saplings measured or estimated for sampling plots. Plot sizes were similar between data sources and all data were converted to per ha values for this analysis. I assume the mixed linear model for trend proposed in Urquhart et al. 1993 is appropriate for analyzing future KLMN vegetation data. I estimate the power to detect a linear trend using this model based on these limited data and the Network's temporal sampling scheme. Based on this, the sampling design seems sufficient to meet the sampling objective of detecting a 50% change over 15 years for basal area, but it may take twice that long to detect 50% change in saplings, which exhibit greater spatial and temporal variation. However, after the power analysis was conducted the Network has modified the protocol to a three year always revisit design. For estimating annual status the proposed change to the revisit design may be negatively affected. However, the switch to an always revisit design should increase the power for detecting trends as compared to the estimated power reported here. While this is encouraging, the current results are provisional; as monitoring data are gathered, power and the model used for analysis should be reassessed.

Introduction

This report summarizes the preliminary univariate power analysis for the vegetation protocol. The power analysis is based on the pilot data provided by Dennis Odion for Basal Area (meters squared/ha) and saplings (per ha) obtained and summarized from US Forest Service Forest Inventory and Analysis plots, and from the pilot study plots undertaken by the Klamath Network at Crater Lake.

Data

I used the data in *KLMNunivariateData revised DCO.xls* for the power analysis. The responses were sapling density (per ha) and plot basal area (m^2/ha). Plot sizes were similar between data sources and all data were converted to per ha values for this analysis. There was only one year of data for the sites at Crater Lake, and two observations for the FIA plots in consecutive years. Thus, it was not possible to pursue an approach that utilizes a mixed linear model for trend with temporal variance components. However, as sampling continues the Network will be able to obtain such estimates and employ the improved methods being developed for trend analysis for panel designs (Starcevich pers. comm.). Here I use the FIA dataset to provide a rough estimate of the residual variance and both the CRLA and FIA datasets to provide estimates of the site-to-site variance.

Revisit Design

The original proposed revisit design for Klamath Network vegetation monitoring is in Table 1. This is the assumed revisit design for the power analysis. Following power analyses, in response to peer-reviews of the protocol, the Network switched to a three year, always revisit design, as shown for panel 1. This change doubles the number of sites revisited at a 3 year frequency. The first year of sampling will not be 2010. The assignment of sites to panels is based on the GRTS sampling design.

	2010	2013	2016	2019	2022	2025	2028	2031	2034	2037	2040
Panel 1	X	X	X	X	X	X	X	X	X	X	X
Panel 2	X										X
Panel 3		X									
Panel 4			X								
Panel 5				X							
Panel 6					X						
Panel 7						X					
Panel 8							X				
Panel 9								X			
Panel 10									X		
Panel 11										X	

Table 1. Proposed Revisit Design for Crater Lake that was used in the power analysis.

Exploratory Data Analysis

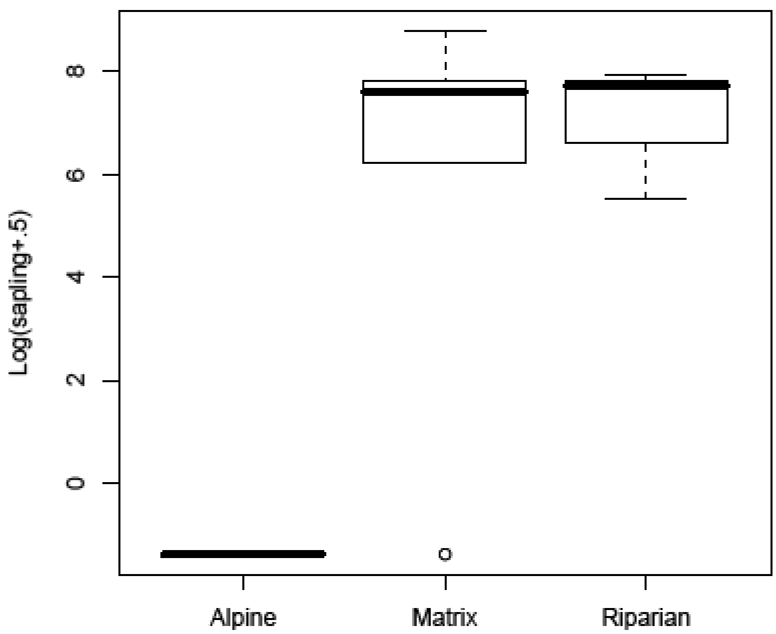

Figure 1. Boxplot of log(sapling density (ha)+.25) for three sampling frames in Crater Lake 2008

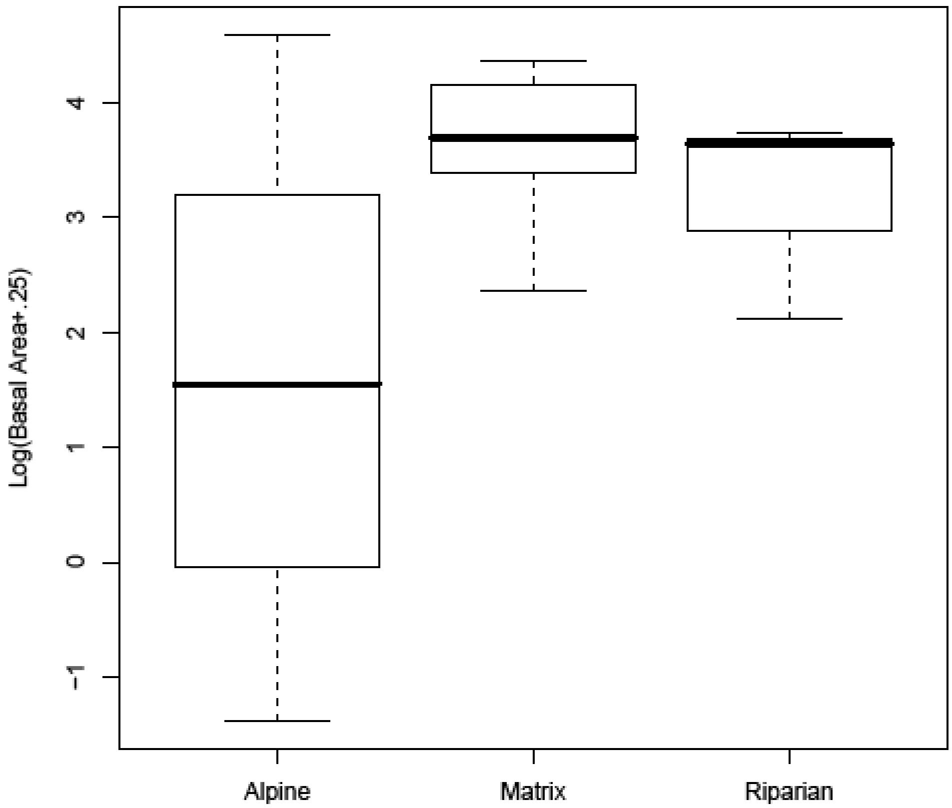

Figure 2. Boxplot of log(Basal Area+.25) for three sampling frames in Crater Lake 2008

There is evidence of one outlier for the Crater Lake matrix sites for sapling density in Figure 1, obviously with such a low sample size it is unclear if this observation is unusual or not. I include the observation for the power analysis. There are no major outliers for Basal Area (Figure 2) in the Crater Lake sites.

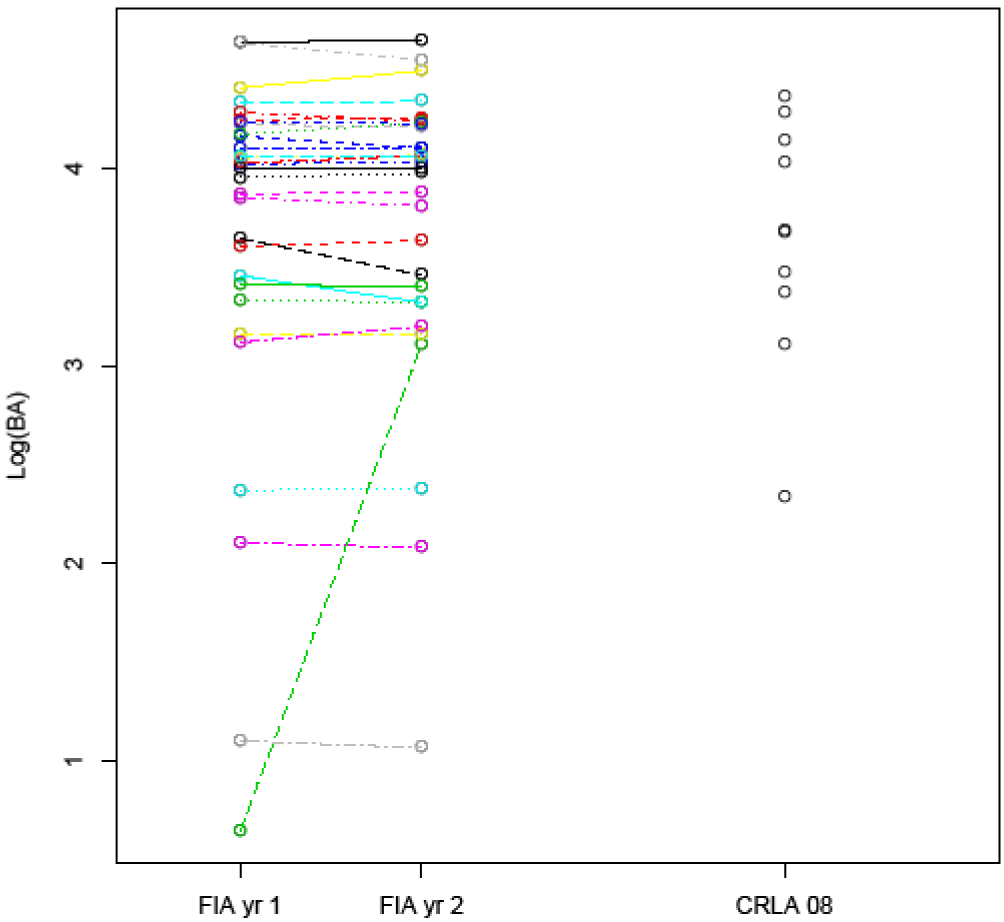

Figure 3. Basal Area for FIA plots and CRLA plots

In the FIA data, plot 5611 is the only one that showed a major increase in log(BA) from year 1 to year 2 in Figure 3. This site is driving the estimate of the residual variance component (site*year). The CRLA matrix sites have less variability among sites compared to the FIA sites.

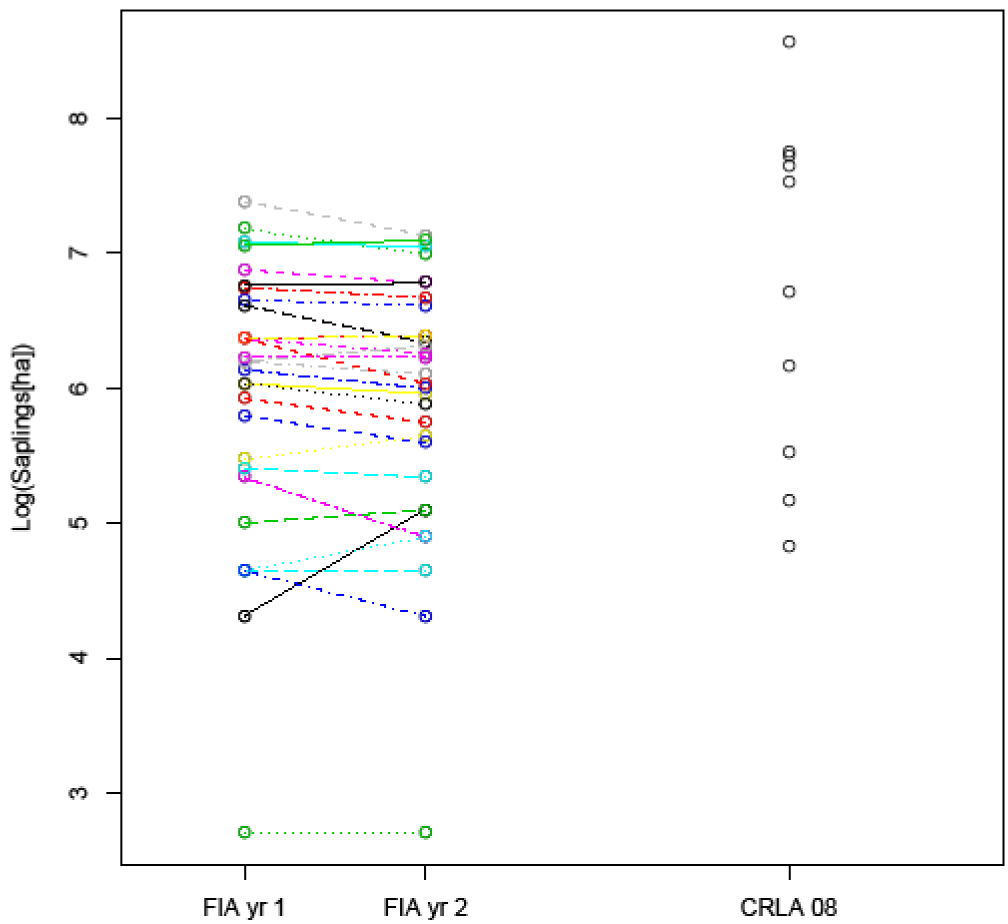

Figure 4. Sapling density for FIA and CRLA plots

Figure 4 displays Sapling density for both the FIA sites and CRLA matrix sites. There is a general pattern of a slight decline in sapling density between the two years FIA data was collected. The CRLA matrix sites appear to have similar site-to-site variability compared to the FIA plots.

Power Analysis

In order to perform a power analysis for univariate trend, a model must be assumed for the future data. I adopt the linear model presented in Urquhart and Kincaid (1999);, Larsen et. al (2001); Kincaid et. al (2004); and Urquhart et. al (1993). The model is as follows where is the observed characteristic of interest (e.g., BA) for site *i* in year *j,* , , ,and the components are assumed independent. There have been many modifications to this general model idea (Piepho and Ogutu, 2002, Van-Leeuwen et al. 1996). I used the functions written by Tom Kincaid to estimate power based on model above, for specific details refer to the paper by Urquhart et al 1993. These are *estimates* of the power because we are estimating the variance components. These estimates will be improved once more sampling is conducted. The model cannot be fully implemented using the Klamath Network's Crater Lake dataset because we do not have estimates for or . Instead we fit a simplified model assuming different values for and based on the pilot datasets from Crater Lake and FIA plots. We use a log transformation such that trend is in terms of a multiplicative change in the medians over time, this is typically appropriate for biological data that display exponential growth and increasing variability with an increase in mean.

Estimated Variance Components

I used the MIXED procedure in the SAS system which can be used to estimate the random and fixed components of mixed models. The estimated variance components using SAS are displayed in Table 2 for Basal Area and Table 3 for Sapling Density.

DATA	**Parameter**	**Estimate**
FIA and CRLA		0.547
		0
		0.093
FIA		0.636
		0
		0.094
CRLA		0.38
		0
		0

Table 2 Estimated Variance Components using REML for Basal Area.

Restricted maximum likelihood (REML) estimates are preferred for unbalanced designs (split panel designs). For the power calculations for Basal Area I assume that , the power is sensitive to the assumption that The estimated site-to-site variability was less for the matrix sites at Crater Lake compared to the FIA sites. To be conservative I will use the estimated variance components for FIA only for Basal Area.

DATA	Parameter	Estimate
FIA and CRLA		1.216
		0.000596
		0.024
FIA		0.964
		0.000539
		0.024
CRLA		1.63
		0
		0

Table 3. Estimated Variance Components using REML for Sapling Density

To be conservative for the power calculations for sapling density, I use the estimated variance components for all observations (FIA and CRLA) for and . Also, I investigate two different values for and assume for sapling density.

Results

For the following power analysis results I investigate four different three-year percent changes in the medians 2.5%, 3.0%, 6.0%, and 10.0%. A 2.5% and 3% per 3-year change correspond to a net change of 25 and 30 percent in the median after 10 sampling occasions (30 years); whereas a 6.0% and 10% per 3-year change corresponds to a net change of 30% and 50% in the median after 5 sampling occasions (15 years).

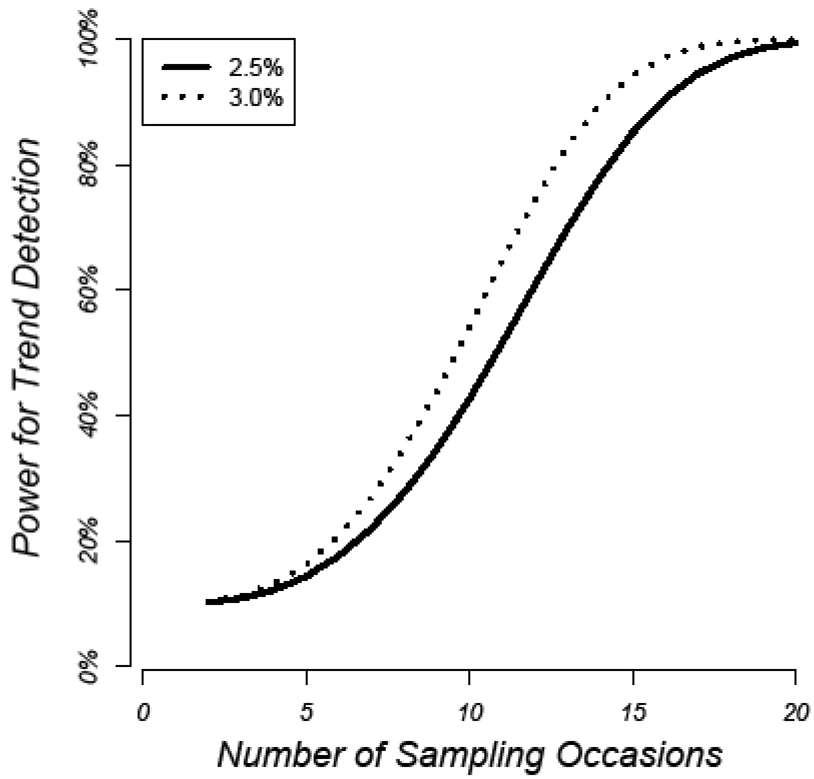

Figure 5. Power for 2.5% and 3.0% three-year trends in median Basal Area.

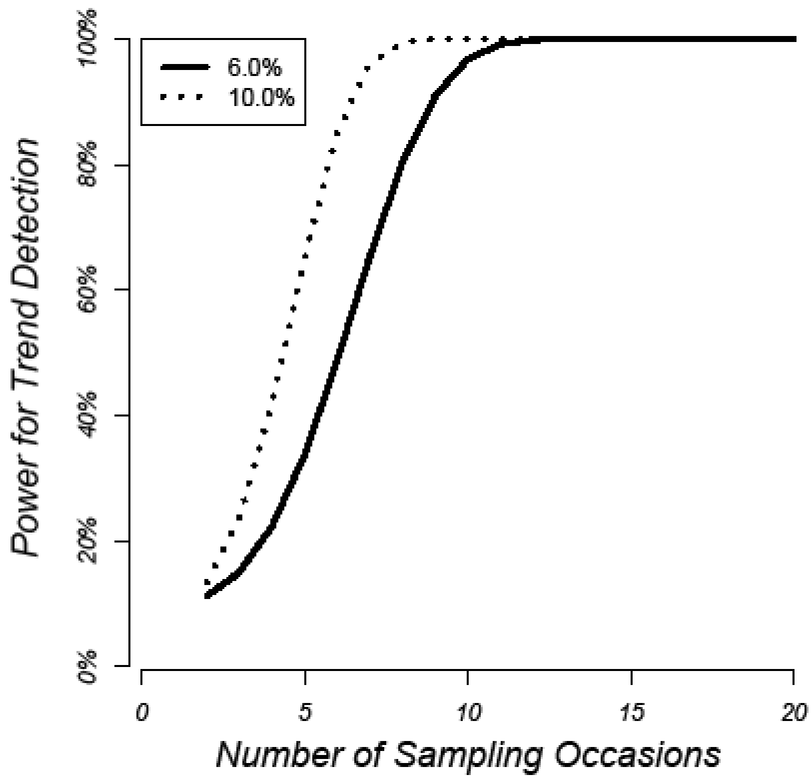

Figure 6. Power for 6% and 10% three-year trends in median Basal Area.

Figure 5 suggests for the proposed revisit design (Table 1) and the given estimated variance components (Table2), the desired 80% power to detect a 2.5% three-year change in median Basal Area will be reached after 14 sampling periods (42 years). For a greater three-year percent change (3.0%) the desired power will be reached after 11 sampling periods (33 years) of every three year sampling for Basal Area. This is an estimate of the power for the chosen sampling design to detect trend, the power will decrease if the true variance components are larger.

The stated objectives of 30% to 50% change over 15 years corresponds to a 6% and 10% three-year percent change, Figure 6 shows that for that magnitude of change it would take 8 sampling occasions or 24 years to detect a 6% three-year trend with 80% power and greater than 15 years (or greater than 5 sampling occasions)to detect a 10% three-year trend in median Basal area with 80% power. Considering that after 5 sampling occasions the total sample size is 90, 6 panels will be visited for one year and 15 sites (1 panel) will be visited every 3-years, the design seems sufficient to meet the sampling objective of detecting a 50% change over 15 years. Again this is assuming that there is no temporal variation or regional variation due to climatic factors or other regional-scale factors, if there is temporal variation the power will decline (see Figure 10). Larsen et. Al 2004 claim that the time variance component cannot be reduced through design choices, but instead through identifying controlling factors (i.e. including regional scale covariates).

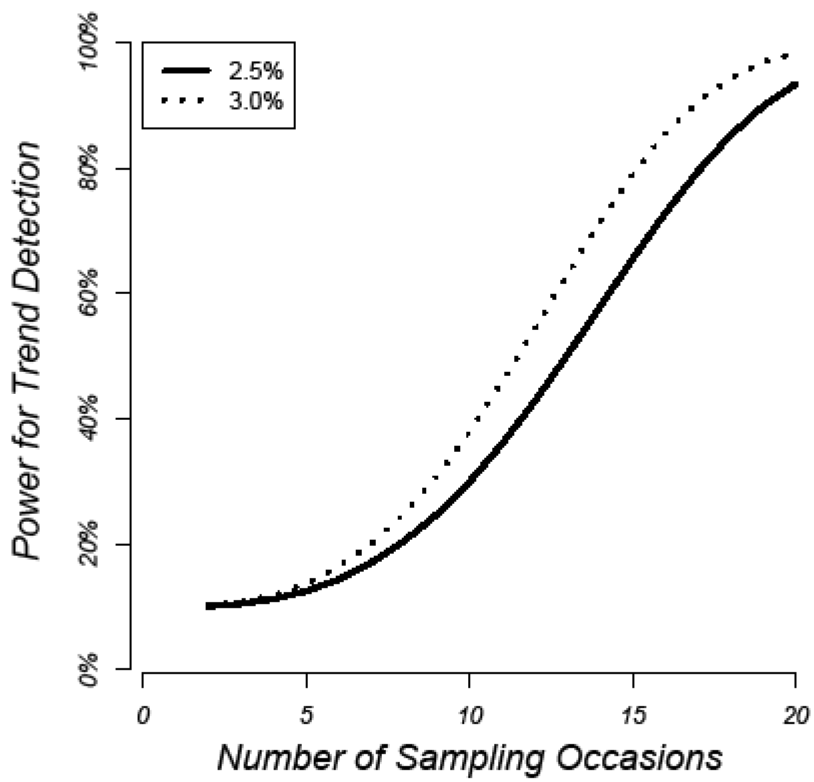

Figure 7. Power for 2.5%and 3.0% three-year trends in Sapling Density .

Figure 7 suggests for the proposed panel revisit design (Table 1) and the given estimated variance components (Table 3) the desired 80% power to detect a 2.5% three-year change in median sapling density will be reached after 17 sampling periods (51 years). For a greater % change (3.0) the desired 80% power to detect trend will be reached after 14 sampling periods (42 years).

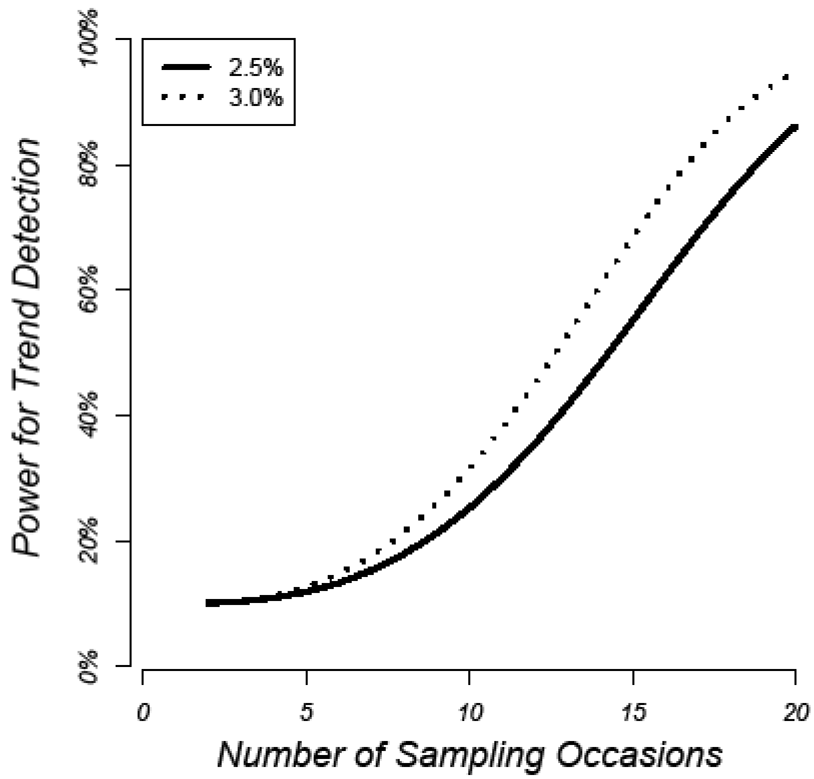

Figure 8. Power for 2.5%and 3.0% three-year trends in Sapling Density .

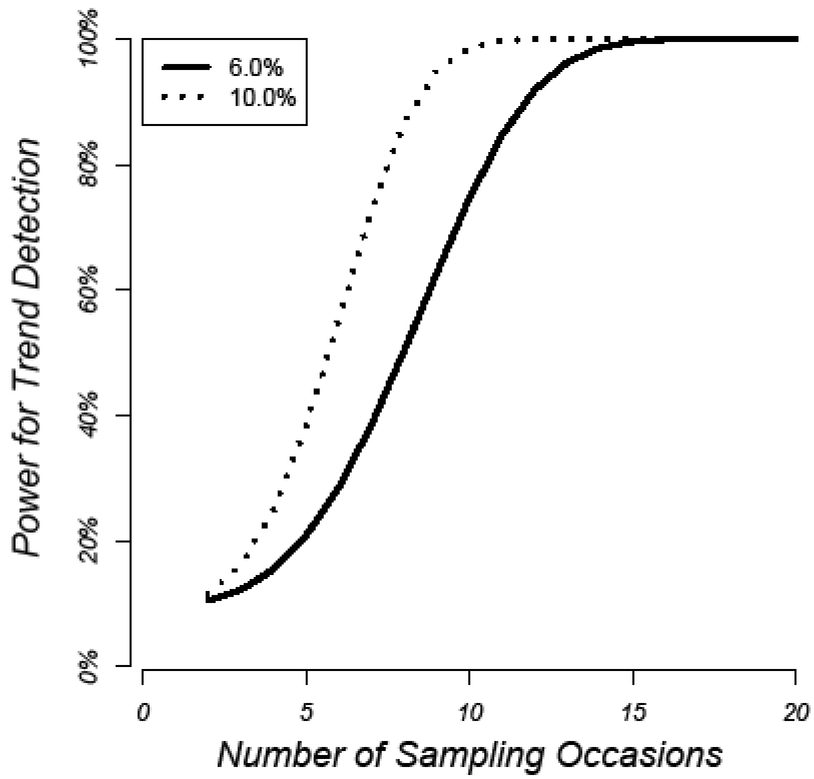

Figure 9. Power for 6.0%and 10.0% three-year trends in Sapling Density .

Figure 8 suggests that with a larger estimated variance component for site-to-site variability the power decreases, as expected. The power to detect a trend is only .70 after 17 sampling periods for a 2.5% three-year trend whereas it is only .61 after 14 sampling periods for a 3.0% three-year trend. The compiled results are presented in Table 4.

Figure 9 shows for a 6% three-year trend in median sapling density it would take more than 30 years to reach the desired power of 80%. For a 10% three-year trend it would take more than 21 years of sampling to detect that level of change in median sapling density with 80% power. It will take longer to reach the desired level of power to detect change in the median sapling density compared to median Basal area primarily because of the non-zero estimate of the variance component for year and the larger estimated variance component for site.

Sampling Occasion (every 3 yrs)	2.50%	1.216 3.00%	2.50%	3.00%
2	0.101	0.102	0.101	0.101
3	0.105	0.107	0.104	0.105
4	0.113	0.118	0.109	0.114
5	0.125	0.136	0.119	0.127
6	0.144	0.163	0.133	0.148
7	0.17	0.2	0.153	0.176
8	0.204	0.248	0.179	0.213
9	0.247	0.308	0.212	0.259
10	0.299	0.378	0.252	0.314
11	0.36	0.458	0.3	0.379
12	0.428	0.543	0.354	0.451
13	0.501	0.63	0.415	0.528
14	0.578	0.714	0.482	0.607
15	0.655	0.79	0.551	0.685
16	0.728	0.855	0.621	0.758
17	0.794	0.906	0.69	0.822
18	0.851	0.943	0.754	0.875
19	0.898	0.968	0.812	0.918
20	0.934	0.984	0.861	0.949

Table 4. Estimated Power for different three-year trends and estimates for Sapling density.

To illustrate the importance of the assumption of , I assume a minimal variance component for Basal area of .05 and the power decreases substantially (Figure 10).

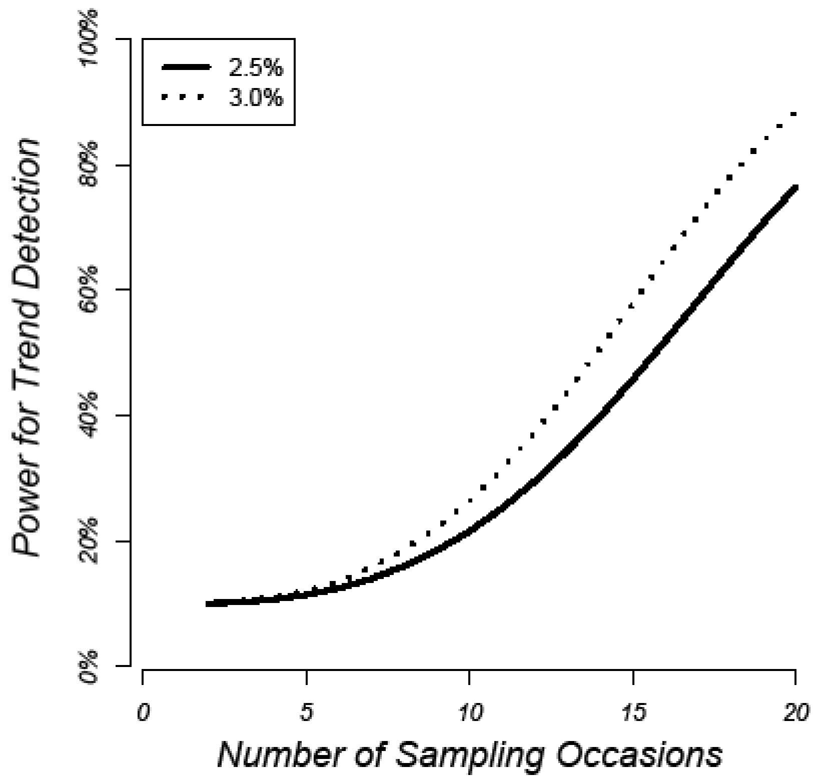

Figure 10. Power for 2.5% and 3.0%three-year trends in Basal Area with .

One point I would like to stress is that this power analysis is based on the assumed model. Starcevich (personal communication) is working on a model for improved trend detection that should be used in the future. At this point it is unknown how the estimated power based on current model versus the revised model will compare.

Appendix:

Example SAS code used to estimate random effects for site and year

```
proc mixed data=FIA method=REML;
  class Site Year;
  model lnBA = ;
  random Site Year;
  run;
```

References

Kincaid, T. K., Larsen, D. P., and Urquhart, N. S. (2004). The structure of variations and its influence on the estimation of status: indicators of condition of lakes in Northeast, U.S.A. Environmental Monitoring and Assessment, 98(12):1–21.

Larsen, D. P., Kaufman, P. R., Kincaid, T. K., and Urquhart, N. S. (2004). Detecting persistent change in habitat of salmon-bearing streams in the Pacific Northwest. Canadian Journal of Fisheries and Aquatic Science, 61:283–291.

Larsen, D. P., Kincaid, T. K., Jacobs, S. E., and Urquhart, N. S. (2001). Designs for evaluating local and regional scale trends. Bioscience, 51(12):1069–1078.

McDonald, T. L. (2003). Review of environmental monitoring methods: survey designs. Environmental Monitoring and Assessment, 85:277–292.

Piepho, H.-P. and Ogutu, J. O. (2002). A simple mixed model for trend analysis in wildlife populations. Journal of Agricultural, Biological, and Environmental Statistics, 7(7):350–360.

Urquhart, N. S. and Kincaid, T. K. (1999). Designs for detecting trend from repeated surveys of ecological resources. Journal of Agricultural, Biological, and Environmental Statistics, 4(4):404–414.

Urquhart, N. S., Overton, W. S., and Birkes, D. S. (1999). Comparing sampling designs for monitoring ecological status and trends: Impact of temporal patterns. Journal of Agricultural, Biological, and Environmental Statistics, 4(4):404–414.

VanLeeuwen, D. M., Murray, L. W., and Urquhart, N. S. (2002). A mixed model for both fixed and random trend components across time. Journal of Agricultural, Biological, and Environmental Statistics, 1(4):435–453.